UNM ALL P9906 Cohort Level 4 Data

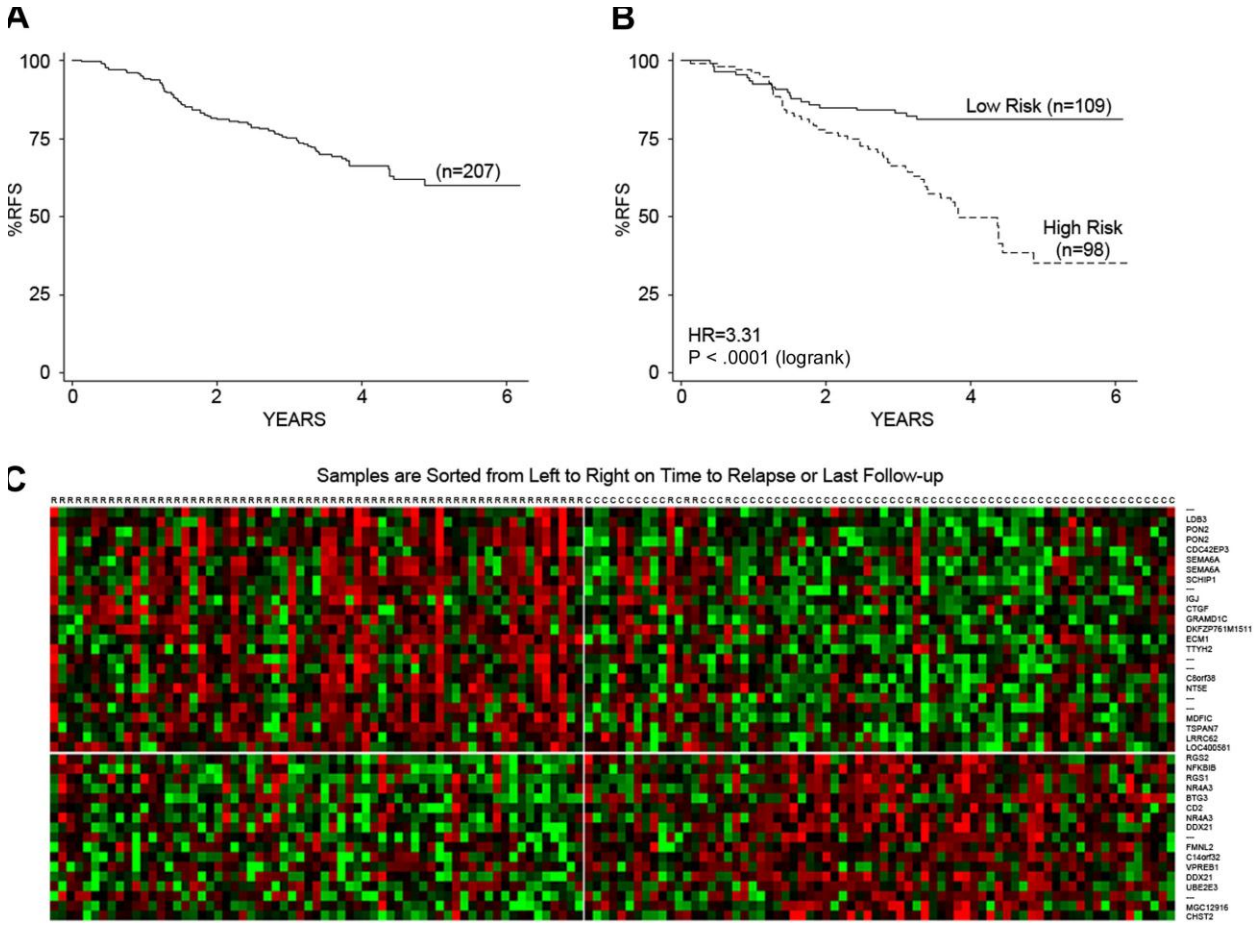
University of New Mexico Bioinformatics Group

27 February 2012

The TARGET collaboration defines ***Level 4 gene expression data*** as a summary of analysis results (case studies and/or published) showing "associations across multiple cases, possibly correlated to clinical data." Level 4 data for P9906 has been compiled and published in the following articles (note: this list is not complete; several additional ms. utilizing the complete cohort or selected subsets are in preparation).

1. *Gene expression classifiers for relapse-free survival and minimal residual disease improve risk classification and outcome prediction in pediatric B-precursor acute lymphoblastic leukemia*. Blood. 2010;115:1394-1405. Huining Kang, I.-Ming Chen, Carla S. Wilson, Edward J. Bedrick, Richard C. Harvey, Susan R. Atlas, Meenakshi Devidas, Charles G. Mullighan, Xuefei Wang, Maurice Murphy, Kerem Ar, Walker Wharton, Michael J. Borowitz, W. Paul Bowman, Deepa Bhojwani, William L. Carroll, Bruce M. Camitta, Gregory H. Reaman, Malcolm A. Smith, James R. Downing, Stephen P. Hunger, and Cheryl L. Willman.

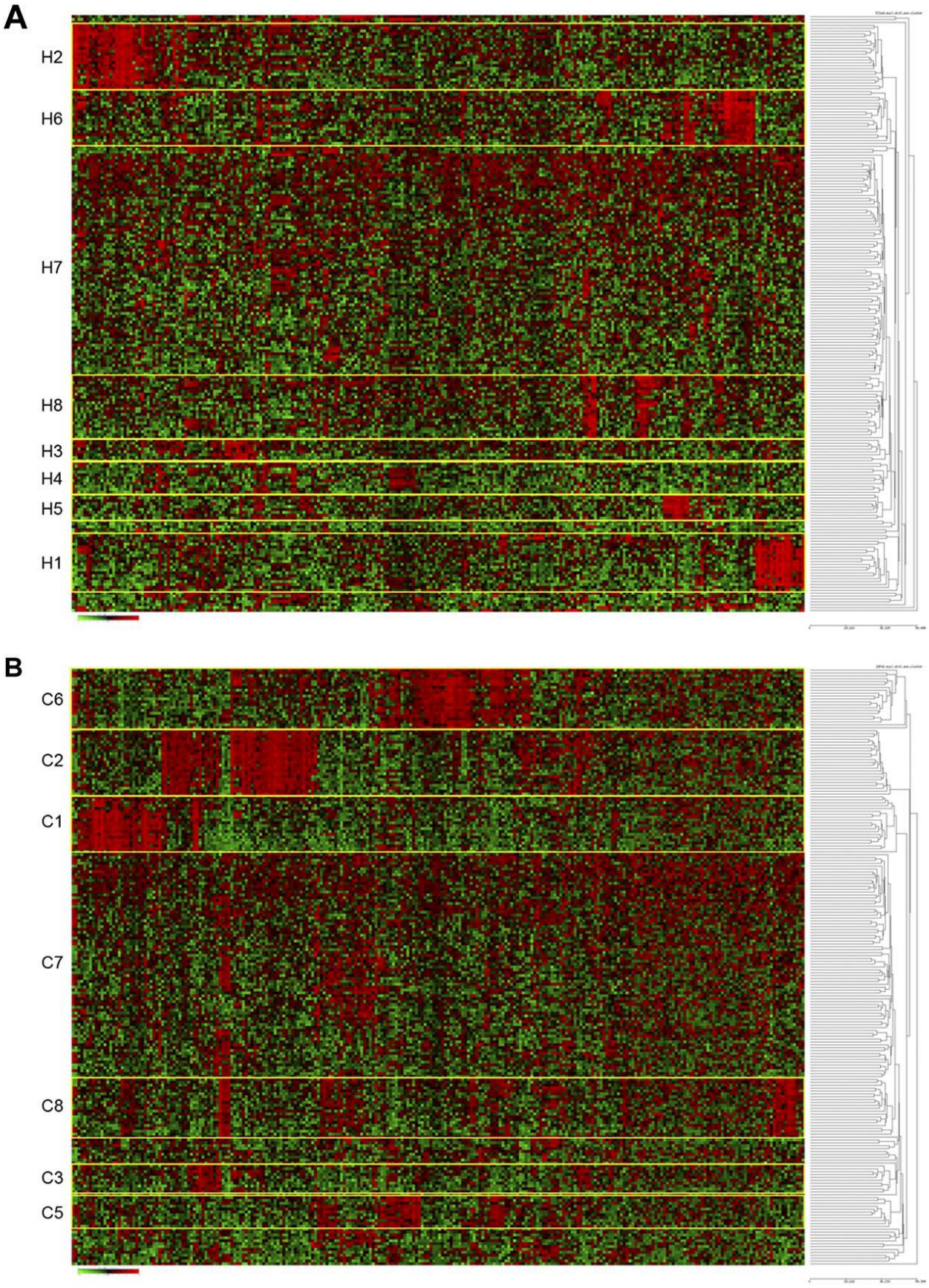
Illustrative figure:
Caption: **Performance of 42-probe-set (38-gene) gene expression classifier for prediction of RFS.** (A-B) Kaplan-Meier survival estimates of RFS in the full cohort of 207 patients (A) and in the low- versus high-risk groups distinguished with the gene expression classifier for RFS (B). HR is the hazard ratio estimated using Cox regression. (C) A gene expression heatmap is shown with the rows representing the 42 probe sets (containing 38 unique genes) composing the gene expression classifier for RFS. The columns represent patient samples sorted from left to right by time to relapse or last follow-up. Red indicates high expression relative to the mean; green, low expression relative to the mean; R, relapse; and C, continuous remission.



2. *Identification of novel cluster groups in pediatric high-risk B-precursor acute lymphoblastic leukemia with gene expression profiling: correlation with genome-wide DNA copy number alterations, clinical characteristics, and outcome.* Blood 2010;116:4874-4884. Richard C. Harvey, Charles G. Mullighan, Xuefei Wang, Kevin K. Dobbin, George S. Davidson, Edward J. Bedrick, I-Ming Chen, Susan R. Atlas, Huining Kang, Kerem Ar, Carla S. Wilson, Walker Wharton, Maurice Murphy, Meenakshi Devidas, Andrew J. Carroll, Michael J. Borowitz, W. Paul Bowman, James R. Downing, Mary Relling, Jun Yang, Deepa Bhojwani, William L. Carroll, Bruce Camitta, Gregory H. Reaman, Malcolm Smith, Stephen P. Hunger and Cheryl L. Willman.

Illustrative figure:

Caption: **Hierarchical clustering identifies 8 cluster groups in high-risk ALL.** Hierarchical clustering using 254 genes (provided in Supplement) was used to identify clusters of patients with shared patterns of gene expression. Rows indicate 207 high-risk ALL patients from COG P9906; and columns, 254 probe sets. Shades of red represent expression levels higher than the median; and green, levels lower than the median. The cluster groups are numbered and prefixed by their method of probe set selection: H indicates high CV; C, COPA; and R, ROSE. (A) HC method for selection of probe sets. (B) COPA selection of probe sets. (C) ROSE selection of probe sets.


A
H2
H6
H7
H8
H3
H4
H5
H1
B
C6
C2
C1
C7
C8
C3
C5
C
R2
R6